

# Transfer of Molecular Property Tensors in Cartesian Coordinates: A New Algorithm for Simulation of Vibrational Spectra

PETR BOUŘ,<sup>1\*</sup> JANA SOPKOVÁ,<sup>1</sup> LUCIE BEDNÁROVÁ,<sup>1</sup>  
PETR MALOŇ,<sup>1</sup> and TIMOTHY A. KEIDERLING<sup>2</sup>

<sup>1</sup>*Institute of Organic Chemistry and Biochemistry, Academy of Sciences of the Czech Republic, Flemingovo nám 2, 16610, Praha 6, Czech Republic; and* <sup>2</sup>*Department of Chemistry, University of Illinois at Chicago, Chicago, IL 60607-7061*

*Received 4 December 1995; accepted 28 July 1996*

## ABSTRACT

A direct transfer of Cartesian molecular force fields (FF) and electric property tensors is tested on model systems and compared to transfer in internal coordinates with an aim to improve simulation of vibrational spectra for larger molecules. This Cartesian transformation can be implemented easily and offers greater flexibility in practical computations. It can be also applied for transfer of anharmonic derivatives. The results for model calculations of the force field and vibrational frequencies for *N*-methylacetamide show that our method removes errors associated with numerical artifacts caused by nonlinearity of the otherwise required Cartesian to internal coordinate transformation. For determination of IR absorption and vibrational circular dichroism intensities, atomic polar and axial tensors were also transferred in the Cartesian representation. For the latter, which are dependent upon the magnetic dipole operator, a distributed origin gauge is used to avoid an origin dependence. Comparison of the results of transferring *ab initio* FF and intensity parameters from an amide dimer fragment onto a tripeptide with those from a conventionally determined tripeptide FF document some limitations of the transfer method and its possible applications in the vibrational spectroscopy. Finally, application to determination of the FF and spectra for helical heptapeptide are presented and compared to experimental results. © 1997 by John Wiley & Sons, Inc.

\* To whom all correspondence should be addressed.

## Introduction

Vibrational spectroscopy (infrared [IR] and Raman) has moved from its traditional base in terms of empirically derived force fields to a larger dependence on *ab initio* quantum-mechanically derived force fields (FF).<sup>1</sup> In the process of developing such FFs, or Hessian matrices of the second derivatives of the molecular energy function, it has become natural to also compute electric dipole derivatives and polarizability derivatives to yield estimates of the IR and Raman intensities. While the FF eigenvalues can be evaluated by comparison to transition frequencies, the comparison of theoretical and experimental intensities provides a good test of the eigenvectors. Put another way, simulation of molecular vibrational spectra (frequencies and intensities) is an excellent overall test of force fields.

Like all *ab initio* quantum-mechanical calculations, FF simulations of molecular spectra have always been limited by molecular size and computer capabilities. While these limitations continually are pushed back by both technological and algorithmic developments, application of *ab initio* FFs to biological molecular systems remains a major problem. A traditional circumvention of the problem for developing FFs for large molecules both at the empirical and *ab initio* levels has been the transfer of parameters from smaller molecules or fragments of a larger molecule to estimate the FF of the large system. The transfer of harmonic force constants in internal coordinates has a long history in vibrational spectroscopy. Vibrational FFs for peptides and nucleic acids have had some success in stimulating the frequency patterns seen experimentally.<sup>2</sup> However, these rarely have the accurate detail needed to obtain high quality simulations of the IR or Raman intensities. In large part this is due to the assembled FF parameters having a fairly good representation of the diagonal, or bond-centered contributions, but a much poorer inclusion of the weaker, longer range interactions. It is these latter effects which lead to the nonuniform dispersion of intensities over "typical" group frequencies that are often used to characterize biopolymer IR and Raman spectra.<sup>2</sup> It is these smaller interactions which tend to vary among different conformational states of the biopolymers and thus it is these that are most important for

using vibrational spectra to carry out conformational analyses.

In recent years, vibrational circular dichroism (VCD) and Raman optical activity (ROA), which are both differential vibrational polarization intensity measurements, have been applied to study a variety of biomolecular conformational problems.<sup>3,4</sup> In small molecule applications of VCD and ROA to stereochemical problems,<sup>5,6</sup> it has been vital to utilize *ab initio* FFs and dipole and polarizability computations (at least at the HF level with a moderate size basis set [e.g., 6-31G\*\*]<sup>7,8</sup>) for the interpretation of observed spectra. Empirical FFs have generally not yet demonstrated enough accuracy to simulate properly the small splitting of near degenerate transitions. While this is not a major problem for IR or Raman intensity calculations, for VCD and ROA, given the signed nature of these differential spectra, misordering of coupled transitions can result in inverting the simulated band shapes in that region of this spectrum which could lead to completely erroneous interpretations.

Originally, internal (local symmetry) coordinates<sup>1,9</sup> were introduced for FF generation with the goal of simplifying the Hamiltonian and providing interpretability in terms of bonding characteristics and promoting ease of transfer to other molecules. Internal coordinates provide a concise representation of normal mode motion, but this advantage is less significant for biopolymeric molecules in which a detailed normal mode assignment is rarely done. Thus direct calculation of the quantum-mechanical FFs in terms of Cartesian coordinates is now the rule. If spectral intensities are to be calculated, a duplicate set of internal coordinates must be defined for the transfer of electric tensors. On the contrary, such a special treatment of the tensors is not required in a Cartesian coordinate representation. Finally, for the anharmonic terms, the requisite nonlinear transformation between internal and Cartesian coordinates is practically impossible, so that maintaining a Cartesian basis throughout is imperative.

The purpose of this study is to demonstrate that the FF and other molecular properties such as dipole derivatives (represented as atomic polar tensors, APT) can be more efficiently transferred from small model systems to larger biomolecular problems of interest in Cartesian coordinates. This approach is predicated upon transfer of atomic properties which can be implemented in a computation for both molecular FF and properties (electric tensors) in a consistent manner.

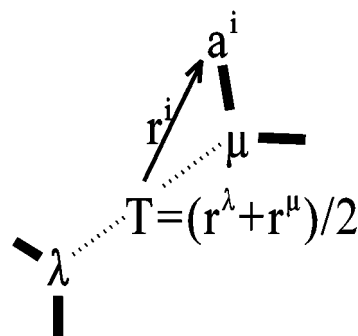
## Method

### FORCE FIELD

The construction of the force field for a larger ("big") molecule typically relies on the chemical and topological similarity between a segment  $F$  of the big system, XXXFXXX, and a part  $f$  of a small molecule or fragment, xfx. For example, if we take the large peptide molecules of prime interest to us acting as the big system, a  $\text{CH}_3\text{—CONH—CH}_2\text{—CONH—CH}_3$  "pseudo-dimer" might provide a suitable small fragment. In this case, if we choose  $F = f = \text{—C—CONH—}$ , the force field associated with the basic amide link can be transferred to the large peptide of interest. Alternatively, for the choice  $F = f = \text{—CONH—C—CONH—}$ , that part of the force field corresponding to the interaction between neighboring amide groups can be transferred as well. In our method, the following steps are carried out to effect the force field transfer:

1. Construct Cartesian coordinates appropriate to describe the nuclear geometry of the big molecule from the best available source. Typically, for a globular protein of known structure the X-ray coordinates can be taken directly or idealized internal coordinates can be transformed into Cartesians for peptides or proteins of unknown structure. (In such a latter case, the spectral analysis will provide one test of the correctness of the assumed structure.)
2. Calculate the *ab initio* force field of the appropriate small fragment (molecule) with a sufficiently large basis set to assure some level of reliability.
3. The crucial step in the Cartesian transfer is to find the best overlap between fragments  $f$  and  $F$  for each atom pair  $\lambda\mu$ , where atoms  $\lambda$  and  $\mu$  are part of the fragment. First, two identical sets of  $m$  atoms,  $\{a^i\}$  and  $\{a^i\}'$  for  $F$  and  $f$  are chosen that contain or are connected to atoms  $\lambda$  or  $\mu$ . Then the orientation of the piece,  $f$ , of the small fragment, xfx, is optimized to maximize overlap with the piece,  $F$ , of the larger molecule, XXXFXXX, by minimization of:

$$\delta(\mathbf{U}) = \sum_{i=1 \dots m} (\mathbf{r}^i(F) - \mathbf{U} \cdot \mathbf{r}^i(f))^2 \quad (1)$$



**FIGURE 1.** Scheme of the Cartesian FF transfer. For the transfer of force constants that include atoms  $\lambda$  and  $\mu$  a coordinate system with origin at  $T$  is defined. The system is then rotated with the aim of overlapping the source and target structures, thus minimizing the RMS distance differences for  $\lambda$ ,  $\mu$ , and connected atoms ( $a^i$ ).

where  $\mathbf{r}^i(F)$  and  $\mathbf{r}^i(f)$  denote coordinates of atoms  $a^i$  with respect to the geometric centers,  $T$  and  $T'$ , respectively, as defined in Figure 1.

In this article, the procedure will be demonstrated by using the pseudo-dimer noted above to describe the vibrational properties of a linear peptide molecule. For example, to obtain a set of force constants related to the oxygen and hydrogen atoms in the same amide group of the peptide ( $\lambda = \text{O}$ ,  $\mu = \text{H}$ ), the  $\text{O}=\text{C—N—H}$  atoms associated with the amide group of the dimer ( $\{a^i\} = \{\text{O}, \text{C}, \text{N}, \text{H}\}$ ) are used as the fragments  $f$  and  $F$ . Note that if the geometries of the  $f$  and  $F$  are the same, one rotation of the coordinate system would be sufficient to allow transfer of the force field for that fragment in a single step.

4. The unitary transformation matrix,  $\mathbf{U}$ , is given by the Euler angles of the rotation and allows transfer of a set of Cartesian force constants for the pair of atoms,  $\lambda\mu$ , from  $f$  to  $F$  by a linear transformation:

$$\begin{aligned} \partial^2 E / (\partial \mathbf{r}_\alpha^\lambda \partial \mathbf{r}_\beta^\mu)(F) \\ = \mathbf{U}_{\alpha\gamma} \mathbf{U}_{\beta\delta} \partial^2 E / (\partial \mathbf{r}_\gamma^\lambda \partial \mathbf{r}_\delta^\mu)(f) \quad (2) \end{aligned}$$

$\alpha, \beta, \gamma, \delta = 1 \dots 3$  (The usual summation convention is used throughout this article; that is, the summation index is that occurring twice in a product.).

If the small molecule contains several fragments,  $f$ , the force constants derived for each can be averaged or weighted according to the distance

between the center of the pair and the center of the set  $\{a^i\}$  for each possible choice. For example, if big = XXXFXXXX, small =  $xfxf'xf''x$ , and  $f = f' = f'' = F$ , three overlaps would be possible and the average of values derived from  $f$ ,  $f'$ , and  $f''$  might provide the most stable transferred force constants. On the other side, some long range interactions in the big molecule may not be encompassed by the small molecule. If an appropriate model is chosen, the latter have negligible effect on the spectra and can be set to zero as a reasonable approximation in many cases. Alternatively, they can be estimated by a calculation on a bigger molecular fragment done at a lower level of approximation, for example, by using CNDO methods or the equivalent.

Cubic and higher anharmonic force constants could also be transferred by this procedure. For cubic constants, three atoms would define the set  $\{a^i\}$  instead of the pair in step 4 (above), and for quartic constants four atoms, etc. We will maintain the harmonic approximation in this introductory study, because of the complexity of the general anharmonic problem.<sup>10,11</sup>

## SPECTRAL INTENSITIES

Derivatives of several electric property tensors with respect to nuclear coordinates are required for calculation of absorption. Raman, VCD, or ROA intensities.<sup>12</sup> These tensors include the molecular electric dipole moment,  $\mu$ , needed for absorption, and the electric dipole–electric dipole polarizability,  $\alpha$ , for Raman intensity. To calculate the optical activity, one has to know also the derivatives of the magnetic moment,  $\mathbf{m}$ , for VCD,<sup>13</sup> the electric quadrupole–electric dipole polarizability,  $\mathbf{A}$ , and the electric dipole–magnetic dipole polarizability,  $\mathbf{G}'$  (“optical activity tensor”), for ROA.<sup>12,14</sup>

While  $\mu$  and  $\alpha$  are independent of translation of the coordinate origin for electrically neutral systems,  $\mathbf{m}$ ,  $\mathbf{G}'$ , and  $\mathbf{A}$  are origin-dependent. This dependence could be a significant impediment to reliable property parameter transfer based on the concept of a *local* chemical similarity. However, this obstacle can be overcome by the introduction of the distributed origin gauge,<sup>12,15</sup> whereby derivatives are expressed in local coordinate systems with their origins on the moving atom:

$$(\partial/\partial x_\varepsilon^\lambda)\mu_\alpha(\lambda) = P_{\varepsilon,\alpha}^\lambda(\lambda) = P_{\varepsilon,\alpha}^\lambda(0) \quad (3a)$$

$$i(2\hbar)^{-1}(\partial/\partial p_\varepsilon^\lambda)m_\alpha(\lambda) = M_{\varepsilon,\alpha}^\lambda(\lambda) = M_{\varepsilon,\alpha}^\lambda(0) \\ - i(4\hbar c)^{-1}\varepsilon_{\alpha\gamma\delta}R_\gamma^\lambda P_{\varepsilon,\delta}^\lambda \quad (3b)$$

$$(\partial/\partial x_\varepsilon^\lambda)\alpha_{\alpha\beta}(\lambda) = (\partial/\partial x_\varepsilon^\lambda)\alpha_{\alpha\beta}(0) \quad (3c)$$

$$(\partial/\partial x_\varepsilon^\lambda)G'_{\alpha\beta}(\lambda) = (\partial/\partial x_\varepsilon^\lambda)G'_{\alpha\beta}(0) + (\omega/2)\varepsilon_{\beta\gamma\delta} \\ \times R_\gamma^\lambda(\partial/\partial x_\varepsilon^\lambda)\alpha_{\alpha\delta}(0) \quad (3d)$$

$$(\partial/\partial x_\varepsilon^\lambda)A_{\alpha,\beta\gamma}(\lambda) = (\partial/\partial x_\varepsilon^\lambda)A_{\alpha,\beta\gamma}(0) \\ - (3/2)\left[R_\beta^\lambda(\partial/\partial x_\varepsilon^\lambda)\alpha_{\alpha\gamma}(0) \right. \\ \left. + R_\gamma^\lambda(\partial/\partial x_\varepsilon^\lambda)\alpha_{\alpha\beta}(0)\right] \\ + R_\delta^\lambda(\partial/\partial x_\varepsilon^\lambda)\alpha_{\alpha\delta}(0)\delta_{\beta\gamma} \quad (3e)$$

In these equations,  $\mathbf{P}$  is the atomic polar and  $\mathbf{M}$  is the atomic axial tensor,  $(\partial/\partial x_\varepsilon^\lambda)$  is the derivative according to the  $\varepsilon$ -coordinate of the atom  $\lambda$ ,  $\mathbf{R}^\lambda$  is the equilibrium position vector of atom  $\lambda$  in the common (e.g., center of mass) origin (0),  $\hbar$  is the Planck constant,  $c$  is the velocity of light, and  $\omega$  is the frequency of the light. The description for the atomic axial tensor [eq. (3b)] used here as a derivative of the magnetic moment with respect to the nuclear momentum (mass-weighted),  $p_\varepsilon^\lambda$ , is consistent with Hamilton’s formalism because  $(\partial/\partial x_\varepsilon^\lambda)m_\alpha = 0$ .

Local tensors defined by eqs. (3) can be transferred using the same transformation matrix,  $\mathbf{U}$ , as used for the diagonal ( $\lambda = \mu$ ) terms of the force field. This apparent simplification results because, in this approximation, the derivatives of the electric tensors are properties of single atoms, while Cartesian force constants are generally related to two atoms. If needed, however, higher order (anharmonic) derivatives could be transferred using the same procedure. Using such anharmonic corrections, the VCD and ROA intensities would be dependent on the origin gauge used for the second derivatives of  $\mathbf{m}$ ,  $\mathbf{A}$ , and  $\mathbf{G}'$ .

Equations useful for simulation of spectral intensities can be found in ref. 16 for VCD and in ref. 17 for ROA. The following formulas for a transition from the ground state  $|0\rangle$  to the first excited state  $|1\rangle$  of a vibrational mode  $i$  were used in this work:

$$\text{Abs: } D_i = |\langle 0|\mu|1\rangle|^2 \\ \langle 0|\mu_\beta|1\rangle = (\hbar/(2\omega_i))^{1/2}P_\beta^{(i)} \quad (4a)$$

$$\text{VCD: } R_i = \text{Im}[\langle 0|\mu|1\rangle \cdot \langle 1|\mathbf{m}|0\rangle] \\ \langle 0|m_\beta|1\rangle = -(2\hbar^3\omega_i)^{1/2}M_\beta^{(i)} \quad (4b)$$

Raman:

$$I_L + I_R = K(7\alpha_{\alpha\beta}^{(i)}\alpha_{\alpha\beta}^{(i)} + \alpha_{\alpha\alpha}^{(i)}\alpha_{\beta\beta}^{(i)}) \quad (4c)$$

(ROA<sub>180</sub>):

$$I_R - I_L = K(8\omega/c) \left[ \omega^{-1} \left( 3\alpha_{\alpha\beta}^{(i)} G_{\alpha\beta}^{(i)} - \alpha_{\alpha\alpha}^{(i)} G_{\beta\beta}^{(i)} \right) + (1/3) \varepsilon_{\alpha\beta\gamma} \alpha_{\alpha\delta}^{(i)} A_{\beta,\gamma\delta}^{(i)} \right] \quad (4d)$$

where **D** and **R** are the dipole and rotational strengths, respectively,  $\omega_i$  is the frequency of the transition, and a derivative according to the normal mode coordinate  $i$  is indicated by superscript ( $i$ ). Such internal coordinate derivatives are related to the Cartesian derivatives from eq. (3) by a linear transformation ( $P_{\beta}^{(i)} = P_{\alpha,\beta}^{\lambda} S_{\alpha,i}^{\lambda}$ , etc., where **S** is the normal mode–Cartesian transformation matrix).<sup>1,9,12</sup> For the ROA parameters in eq. (4d) (selected for the 180° backscattering geometry), intensities for the left ( $I_L$ ) and right ( $I_R$ ) polarized light are specified and  $\omega$  is the frequency of the exciting light.

To carry out the computations in this study, the CADPAC program package<sup>18</sup> proved to be a convenient tool, because it provides a means of generating all of the tensors mentioned above. The derivatives of **μ** and **m** were obtained analytically at the HF level for computation of IR absorption and VCD intensities. For Raman spectra, **α** can be similarly obtained, but derivatives of **A** and  $\omega^{-1}\mathbf{G}'$  must be calculated by a finite difference method.<sup>19,20</sup>

## Applications

### AB INITIO CALCULATIONS OF FORCE FIELDS

Molecular force fields and other properties are mostly computed using equilibrium geometries, although approaches accepting a nonzero gradient of energy are known.<sup>21</sup> Thus, evaluations of different force fields often demand generation of geometries by various *ab initio* calculations performed at different levels. To facilitate analysis, transfer of one force field onto each geometry could be used to separate the error caused by the variations in geometry from errors caused by other sources. As an example, we have chosen to focus on *N*-methylacetamide (NMA), the simplest amide directly useful as a “small” molecule or fragment for peptide studies with our transfer methods. Here we compare the results of a relatively “high” level of

calculation, particularly useful for vibrational frequencies, using the Becke 3LYP DFT functional<sup>22</sup> and the 6-31G\*\* basis set on its equilibrium-optimized geometry with transfer of this FF to a geometry from a simpler level of calculation, HF/6-31G\*\* geometry, and to an X-ray determined geometry,<sup>23</sup> with both the Cartesian and internal methods of transfer. The GAUSSIAN 94 program package<sup>24</sup> was used for the *ab initio* calculations. Table I compares equilibrium geometry parameters for DFT, HF, and X-ray structures. As can be seen the bond lengths and the bond angles vary within 1–6% for the three geometries. For this example, the coordinates of the hydrogen atoms (poorly located by the X-ray method<sup>25</sup>) were considered same for the X-ray structure as for the DFT geometry.

Calculated harmonic frequencies are given in Table II for the B3LYP/6-31G\*\* FF ( $\omega_1$ ) and HF/6-31G\*\* FF ( $\omega_2$ ), both calculated in their corresponding equilibrium geometries; the B3LYP transferred onto the HF geometry with Cartesian ( $\omega_3$ ) and internal ( $\omega_4$ ) coordinate transfer; and finally, onto the X-ray structure with Cartesian ( $\omega_5$ ) and internal ( $\omega_6$ ) transfer. The DFT frequencies are mostly lower than the HF values and closer to the experimental results.<sup>26,27</sup> However, for the purposes of this test method, we take the B3LYP/6-31G\*\* DFT results ( $\omega_1$ ) to be the reference set and calculate errors with respect to these values. The errors were obtained via the fit  $\omega_j = a\omega_1$ , where  $(a - 1)$  is the systematic error and  $j = 3, 4, 5, 6$ .

**TABLE I.** Main Geometry Parameters for the Three Structures of *N*-Methylacetamide.

Parameter	B3LYP/6-31G**	HF/6-31G**	X-ray
$d(\text{C}—\text{C})$ [Å]	1.523	1.515	1.536
$d(\text{CO}—\text{N})$	1.370	1.355	1.290
$d(\text{N}—\text{C})$	1.451	1.445	1.465
$d(\text{C}=\text{O})$	1.224	1.200	1.236
$d(\text{N}—\text{H})$	1.008	0.993	(1.008) <sup>a</sup>
$\angle(\text{C}—\text{C}=\text{O})$ [deg.]	122.7	122.2	120.5
$\angle(\text{C}—\text{C}—\text{N})$	114.5	114.6	116.5
$\angle(\text{C}—\text{N}—\text{C})$	122.6	123.2	120.5

<sup>a</sup> For the X-ray structure the B3LYP parameters were used for the hydrogen coordinates. For all structures,  $C_s$  symmetry considered; torsion angles  $\angle(\text{H}—\text{C}—\text{C}=\text{O}) = \angle(\text{O}=\text{C}—\text{N}—\text{C}) = \angle(\text{C}—\text{N}—\text{C}—\text{H}) = 0^\circ$ .

**TABLE II.** Tests of Transfer Methods for *Ab Initio* Becke 3LYP / 6-31G\*\* FF of *N*-methylcetamide onto Alternate Geometries.

Mode	B3LYP $\omega_1$	HF $\omega_2$	HF geometry		X-ray geometry	
			$\omega_3$	$\omega_4$	$\omega_5$	$\omega_6$
1	3642.83	3899.49	3642.55	3642.77	3643.92	3642.95
2	3178.89	3333.65	3179.16	3179.24	3178.84	3178.90
3	3167.65	3320.93	3166.89	3167.45	3167.07	3167.67
4	3115.48	3255.69	3116.37	3115.76	3115.71	3115.48
5	3089.87	3239.36	3089.59	3089.88	3090.17	3089.78
6	3051.95	3194.20	3052.17	3051.17	3051.83	3051.97
7	3036.13	3187.64	3036.08	3036.50	3036.23	3036.12
8	1791.45	1963.37	1791.53	1794.12	1791.72	1796.51
9	1570.68	1721.72	1569.64	1582.15	1571.71	1573.97
10	1515.98	1635.82	1516.89	1534.13	1515.67	1515.94
11	1510.04	1632.80	1509.89	1513.92	1509.23	1510.90
12	1499.69	1612.90	1500.12	1507.12	1498.84	1499.83
13	1484.83	1597.59	1485.01	1493.14	1485.63	1485.16
14	1449.89	1582.83	1450.09	1468.46	1450.53	1449.75
15	1403.54	1534.36	1402.70	1416.31	1402.70	1403.64
16	1273.57	1390.59	1274.42	1282.92	1271.25	1270.39
17	1173.34	1280.05	1173.04	1185.08	1175.19	1178.88
18	1157.58	1255.69	1157.89	1161.16	1157.93	1155.85
19	1117.63	1202.40	1118.09	1121.83	1116.34	1113.85
20	1052.85	1158.18	1053.92	1063.27	1054.20	1057.95
21	975.37	1060.50	975.48	981.24	979.83	980.36
22	874.08	941.60	874.00	876.98	871.94	882.08
23	630.27	681.50	631.51	636.54	628.80	632.93
24	618.92	672.80	618.13	631.73	621.64	630.85
25	436.16	475.30	435.96	441.60	434.19	451.62
26	431.78	463.90	430.83	436.33	433.53	434.09
27	287.68	307.17	288.88	290.12	284.37	285.68
28	158.17	161.15	157.63	160.33	157.48	153.04
29	84.29	<i>i</i> 36.83	99.73	85.50	66.11	84.27
30	<i>i</i> 86.60	<i>i</i> 63.06	<i>i</i> 102.72	.00	<i>i</i> 80.36	.00
Systematic error [%]			0.005	0.2	0.003	0.05
RMS dispersion [%]			4.1	7.0	3.8	4.6

Frequencies [per centimeter]:  $\omega_1$  and  $\omega_2$ , the Becke 3LYP / 6-31G\*\* and HF / 6-31G\*\* *ab initio* calculations, respectively, the former corresponds to the model FF;  $\omega_3$  and  $\omega_4$ , frequencies obtained after the model FF is transferred to the HF / 6-31G\*\*-optimized geometry in Cartesian and internal coordinates, respectively;  $\omega_5$  and  $\omega_6$ , transfer of the model FF onto the X-ray geometry in Cartesian and internal coordinates, respectively.  $i = \sqrt{-1}$ .  $\omega_4$  and  $\omega_6$  frequencies for the 30th mode are not included in the statistics (imaginary frequencies not given by the programs handling the internal coordinate FF).

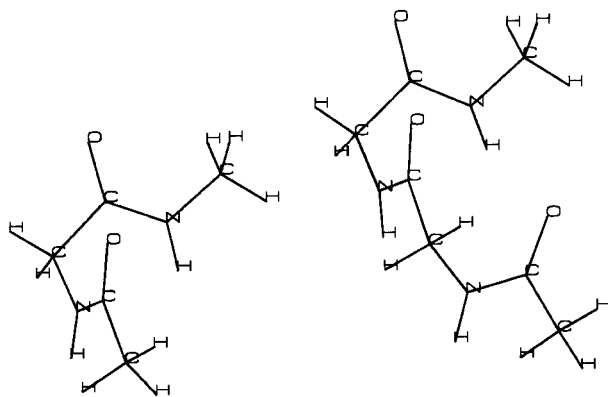
For both the HF and X-ray structures of NMA, if the FF is transferred in internal coordinates instead of the Cartesian coordinates the systematic error increases by more than an order of magnitude. Similarly, the RMS deviation rises significantly for the internal transfer. Presumably, the difference is caused by neglecting nonlinear terms in the Cartesian to internal FF transformation which means the internal transfer has lost some part of the original FF. For C—H, C—D, and C=O stretching modes, the two methods give

very similar frequencies. However, the internal transfer is much less accurate for the lower frequency, deformation modes, which often involve more mixing of local coordinates, where deviations as high as 5–20 cm<sup>-1</sup> can appear. These differences would contribute substantially to the overall error of the transferred DFT calculation with respect to the experiment. By contrast, the Cartesian transfer faithfully reproduces frequencies throughout the entire frequency region (compare  $\omega_1$  and  $\omega_3$ ).

Although the X-ray geometry differs even more from the DFT structure than the HF geometry (except for the hydrogen coordinates; see Table I), only minor changes of frequencies appear if the DFT Cartesian FF is transferred onto the X-ray geometry. The overall errors are even smaller than those found for the case above of transfer onto the HF geometry. These results strongly support our contention that the transfer of FF in Cartesian coordinates minimizes the effect of geometry variations on calculated frequencies. With that principle having been established, we can move on to a more realistic test.

### CONSTRUCTION OF FF FROM SMALLER FRAGMENTS

To test if our idea of using part of the FF of a small species would give reasonable FF for a "big" molecule, we carried out a test using the diamide fragment FF (such as the one we have developed earlier for the  $\text{CH}_3\text{—CONH—CH}_2\text{—CONH—CH}_3$  pseudo-dimer<sup>28</sup>), to generate a FF for the analogous  $\text{CH}_3\text{—CONH—CH}_2\text{—CONH—CH}_2\text{—CONH—CH}_3$  "pseudo-trimer." Then we shall compare the transferred trimer result with that obtained with a more conventional *ab initio* determination of the trimer FF and intensity parameters. For these model calculations, we also have transferred electric and magnetic transition dipole parameters expressed in terms of atomic polar (APT) and axial (AAT) tensors, respectively, and compared them to the directly computed trimer values. To select a geometry, the tripeptide and dipeptide were both constrained to have chain torsional angles as would be found in an  $\alpha$ -helical structure ( $\omega, \phi, \psi = 180^\circ, -47^\circ, -57^\circ$ , respectively) and then minimized. The resulting geometries are shown in Figure 2. The FF and APT parameters for both trimer and dimer were calculated with GAUSSIAN 94 using the B3LYP/6-31G\*\* method as employed above for *N*-methylacetamide. It was not possible to compute the AAT with the DFT method; therefore, following the example of Stephens and coworkers<sup>29</sup> we computed them at the HF level using a 6-31G basis set and the CADPAC program. The results were tabulated (Table III) as frequencies, dipole and rotational strengths from the direct computation of the trimer ( $\omega_1, D_1, R_1$ , respectively), and from the internal ( $\omega_2, D_2, R_2$ ) and Cartesian ( $\omega_3, D_3, R_3$ )



**FIGURE 2.** Geometries of model di- and tripeptides used for the transfer test. The geometries were optimized at the Becke 3LYP/6-31G\*\* level with the torsion angles constrained to mimic the  $\alpha$ -helical conformation.

transfers of FF, APT, and AAT from the dipeptide calculation.

Unlike the previously discussed NMA FF, we cannot transfer all the force constants for the tripeptide. In other words, some long range force constants are arbitrarily set to zero for this test because they do not exist in the "small" dimer used for transfer to the "big" trimer. Even the transferred force constants corresponding to the interaction of atoms in neighboring amide groups, for example, differ in the di- and tripeptide. Thus, the end result of the two methods must be different at some level. Hence, a comparison of the qualities of the internal and Cartesian transfer based solely on frequency precision will not adequately discriminate between the two transfer methods. Rather we compare the spectral parameters obtained by the transfer with those from the "classical" *ab initio* calculation to see if we can obtain a reasonable simulation of the "correct" result, here assumed to be the latter direct calculation. Nevertheless, it should be noted that the superiority of the Cartesian transfer is in its universality and greatest simplicity, which is maintained here without loss of accuracy as compared to the more traditional transfer in internal coordinates, which in this case had to be specially developed and optimized for peptides.<sup>30</sup>

Comparing the transfer methods to the direct calculation, the average frequency error is comparable for both, while the systematic error is smaller for Cartesian transfer. However, these differences are negligible for most frequencies with respect to a typical experimental resolution. Thus, the effect

of inaccurate and “left out” force constants probably combines with and overwhelms the contribution of the nonlinearity error in internal transfer.

The same APTs and AATs were used for the two transferred calculations shown in Table III (columns 6, 7, 9, and 10); however, the two sets of spectral intensities ( $D_2$ ,  $D_3$ ) from the transferred

atomic parameters deviate much more from the directly calculated trimer result ( $D_1$ ) than do the frequencies. This is in accord with the generally known fact that eigenvectors (of the GF matrix<sup>1</sup>) are more sensitive to force field changes than are the eigenvalues (frequencies). Also, the sole dipolar derivatives are can usually be calculated with a

**TABLE III.**  
FF, APT, and AAT Transfer of Dimer Parameters onto a Model Tripeptide. Comparison of Direct Tripeptide *Ab Initio* Calculation Result with that from Transfer in Internal and Cartesian Coordinates from the Dipeptide.

Mode	<i>Ab initio</i>			In internals			In Cartesians		
	$\omega_1$	$D_1$	$R_1$	$\omega_2$	$D_2$	$R_2$	$\omega_3$	$D_3$	$R_3$
78	3636.6	1.7	6.7	3643.0	2.0	6.1	3640.5	1.9	2.0
77	3631.0	2.7	-9.4	3634.5	1.4	19.1	3639.8	2.2	-5.1
76	3582.3	8.6	-20.1	3625.7	3.5	-28.7	3601.9	2.3	-1.7
75	3178.1	0.5	-0.5	3175.0	0.6	-1.3	3175.4	0.6	-0.5
74	3152.2	2.3	-1.5	3147.9	2.3	5.6	3148.5	2.3	-1.1
73	3139.0	0.7	5.7	3145.3	1.1	-1.1	3137.1	0.8	6.6
72	3127.7	0.7	6.9	3126.5	0.9	10.4	3132.4	1.4	6.3
71	3118.7	3.2	-5.2	3115.7	1.7	-4.2	3115.6	1.5	-4.3
70	3117.6	1.1	3.9	3113.9	2.8	2.6	3113.7	2.9	2.3
69	3082.2	2.4	1.7	3079.6	2.3	3.5	2079.3	2.7	-4.1
68	3073.9	2.1	2.8	3071.7	2.5	3.8	3078.4	2.3	10.1
67	3053.3	0.7	3.4	3051.7	0.9	2.5	3051.4	1.0	3.7
66	3048.8	4.9	-7.8	3049.9	5.3	-7.3	-3049.1	5.1	-7.8
65	1809.4	70.3	-342.2	1806.4	92.0	-487.5	1807.4	90.3	-490.9
64	1795.9	24.2	186.9	1800.3	29.1	188.9	1799.8	28.1	173.9
63	1774.9	44.9	99.2	1787.3	12.5	267.9	1788.8	14.4	278.8
62	1552.3	61.0	-30.3	1549.6	44.4	-12.0	1551.9	51.5	-136.1
61	1543.4	55.8	-106.4	1538.2	49.6	-18.7	1540.2	50.1	110.0
60	1541.5	9.8	31.7	1533.8	47.5	-55.2	1536.6	42.8	-30.2
59	1516.1	2.3	3.0	1519.3	8.4	5.6	1520.5	5.4	3.6
58	1497.1	2.5	-0.0	1498.5	1.6	0.5	1498.0	1.3	-3.2
57	1492.4	4.1	20.1	1490.2	2.1	8.2	1490.6	2.9	-25.4
56	1486.1	3.1	15.0	1482.5	10.2	1.2	1484.3	6.4	26.7
55	1478.6	9.6	-46.5	1476.9	4.2	-41.8	1480.4	2.4	-23.5
54	1474.9	1.6	-40.0	1463.8	4.9	-8.5	1474.1	4.0	-45.8
53	1462.8	1.9	-11.3	1462.0	2.3	-21.4	1464.4	6.7	-16.5
52	1401.9	6.1	5.3	1400.1	5.1	4.6	1399.8	5.1	1.8
51	1361.1	4.6	14.2	1360.8	10.9	26.6	1374.7	7.1	16.3
50	1353.4	8.1	28.6	1311.9	3.2	54.9	1344.8	8.9	27.4
49	1319.0	4.4	26.5	1297.5	0.4	-1.7	1311.6	4.5	27.6
48	1297.3	13.2	7.7	1279.5	22.7	42.1	1282.9	18.3	36.4
47	1273.6	25.3	-33.6	1263.8	26.9	-75.9	1264.5	23.1	-0.8
46	1261.7	12.1	94.0	1250.7	10.1	129.0	1254.7	10.4	61.8
45	1241.6	16.7	70.7	1220.7	12.4	25.5	1223.4	14.0	29.6
44	1187.8	0.1	0.3	1191.3	0.5	3.2	1188.3	1.0	8.5
43	1149.1	1.6	0.4	1155.4	0.7	-3.8	1153.5	0.3	-1.1
42	1117.6	2.0	0.9	1123.6	1.6	1.8	1120.3	1.4	-0.4
41	1112.9	1.6	-15.9	1108.6	2.3	-18.8	1107.6	1.8	-14.9
40	1093.6	1.8	-19.7	1106.0	1.5	-11.5	1097.7	1.1	-12.1



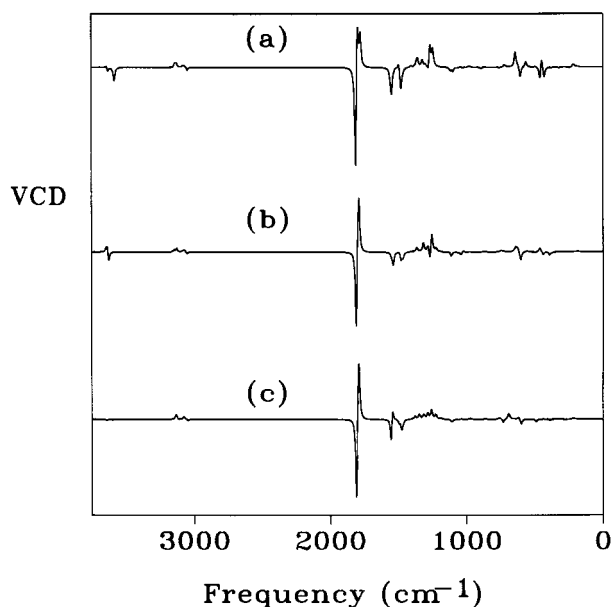
TABLE III.  
(continued)

Mode	Ab initio			In internals			In Cartesians		
	$\omega_1$	$D_1$	$R_1$	$\omega_2$	$D_2$	$R_2$	$\omega_3$	$D_3$	$R_3$
39	1054.9	1.9	−4.1	1054.3	2.0	−6.7	1054.1	2.1	−4.2
38	1035.7	0.7	4.3	1034.9	4.8	−21.0	1032.9	1.3	4.1
37	1026.0	2.6	−5.3	1017.6	0.8	9.0	1024.2	2.7	−2.1
36	974.4	2.6	6.5	971.8	3.5	1.2	970.9	4.0	5.4
35	887.8	2.8	−10.0	880.9	1.1	0.6	877.4	2.0	−5.1
34	872.1	0.4	0.4	861.3	2.2	7.2	857.5	1.2	8.2
33	859.9	3.6	1.2	860.7	2.4	−2.7	847.4	3.6	−1.9
32	749.4	3.1	−7.3	741.5	3.7	15.7	752.2	1.8	16.6
31	715.7	7.7	17.8	719.0	5.2	5.3	726.3	1.9	−37.1
30	648.7	2.1	−0.6	647.5	1.4	12.1	686.3	50.7	69.4
29	636.0	27.9	119.5	633.8	12.4	56.9	646.1	1.1	5.5
28	616.2	2.0	16.2	621.1	5.1	50.3	615.3	1.0	27.5
27	600.0	10.4	−89.6	598.3	14.4	−118.4	598.1	4.2	18.2
26	559.5	27.5	49.8	550.2	30.0	13.7	594.7	6.3	−89.8
25	489.8	13.2	9.5	490.7	15.0	−19.1	485.9	17.9	−47.5
24	454.8	26.8	−150.4	458.1	33.3	70.5	451.3	16.7	−16.8
23	443.5	81.9	136.5	433.6	59.4	−58.8	445.3	81.6	13.1
22	425.0	48.7	−122.4	387.2	72.8	−71.1	382.4	70.2	−31.9
21	372.6	1.8	−6.3	369.4	2.8	−17.8	368.6	2.5	3.6
20	318.2	17.3	−13.5	324.6	23.5	−2.3	323.1	15.0	10.2
19	275.7	23.6	4.8	311.1	13.3	−0.9	272.0	28.1	−15.2
18	240.4	11.9	−15.6	235.6	11.9	−8.7	244.7	4.9	−4.7
17	211.4	23.5	57.5	217.0	14.2	12.8	212.3	17.3	48.7
16	179.6	10.5	21.1	182.0	17.8	32.5	177.8	11.2	18.3
15	170.1	1.4	−6.0	165.3	9.2	11.0	153.0	16.7	−1.7
14	99.9	4.9	−3.0	95.9	19.1	0.6	109.1	4.1	3.1
13	97.8	5.2	8.5	82.4	8.6	−18.6	73.9	9.7	−1.0
12	76.1	35.0	0.1	69.4	65.0	12.4	48.2	71.2	−0.5
11	75.0	21.4	10.0	59.4	27.9	−5.3	35.0	14.4	4.7
Systematic deviation [%]				−0.03	−2	−8	0.015	−0.4	16
RMS dispersion				12	10	51	12	11	56

$\omega_1$  [cm<sup>−1</sup>],  $D_1$  [10<sup>−3</sup> debye<sup>2</sup>], and  $R_1$  [10<sup>−8</sup> debye<sup>2</sup>] frequency, dipole strength, and rotatory strength, respectively, calculated *ab initio*; similarly,  $\omega_2, D_2, R_2$  ( $\omega_3, D_3, R_3$ ): values obtained from the dimer with internal (Cartesian) transfer. RMS dispersions given in units of  $\omega$ ,  $D$ , or  $R$ . Frequencies 1–10 (omitted) are the imaginary and space motion modes.

lower precision than the FF if the same *ab initio* level is used. The VCD intensities show an even greater sensitivity to such modeling than do the absorption parameters. However, in both the calculated spectral intensities ( $D$  and  $R$ ) the overall character of the spectrum, mainly the VCD signs and relative intensities for the most intense bands, are well reproduced, as can be seen in Figure 3. Nevertheless, for some of the C—H stretches, it seems that the ordering of modes has changed after the transfer, such as with the nearly degenerate 3118.7 and 3117.6 cm<sup>−1</sup> modes, which reverse their relative dipole strengths for the transfer cal-

culations. Similarly the C=O stretches (amide I) have a partial intensity reversal, indicating a change in mixing, between the 1795 and 1774 cm<sup>−1</sup> modes, as do the N—H deformation (amide II) modes at 1543 and 1541 cm<sup>−1</sup>. This affects the apparent splitting and intensity distribution in the VCD, in particular, for what would be the amide I and II bands. More severely affected is the amide III region which is known to be affected by NH—C<sub>α</sub>H mixing.<sup>31</sup> Because these model calculations are effectively for glycyl oligomers having two α—Hs per residue, it is difficult to relate such calculations to experimental observations.



**FIGURE 3.** Simulated VCD spectra of the model tripeptide. The *ab initio* calculation (a). The spectrum obtained via the transfer of the FF from the dipeptide in internal (b) and Cartesian (c) coordinates.

Nonetheless, with respect to possible future applications of such a modeling it is gratifying that the calculations all show the qualitative features of a typical  $\alpha$ -helical VCD pattern in terms of net sign of the approximately degenerate amide modes.<sup>32,33</sup> This is best seen by comparison of simulated VCD spectra as shown in Figure 3. For the amide A mode (N—H stretch) the VCD is computed to be a negative couplet (negative then positive with increasing wave number), as seen experimentally in model polypeptides,<sup>34</sup> but the net intensity (summed over the three modes) is negative rather than the typically observed positive. The amide I is computed to be a positive couplet with more intensity in the negative lobe and the amide II is net negative as seen experimentally for all  $\alpha$ -helices.<sup>32,33,35</sup> The amide III is a mixed state that lies between 1250 and 1300  $\text{cm}^{-1}$ , which makes comparison to experiment difficult. In these computations, the net intensity is positive as seen in some model helices.<sup>36</sup> Thus, we see that fundamental and transferred calculations both give the type of spectral interpretation we desire for a tripeptide (assuming it maps onto a larger peptide of defined structure), but that, in the fine details, the transfer has an effect on the FF that causes a change in the detailed mode character.

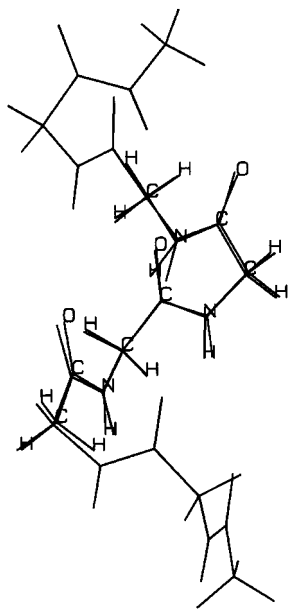
## COMPLEX SIMULATIONS OF INFRARED SPECTRA

To see if the size of the molecules is of importance and to compare more directly to experimental observations, we next simulated the VCD spectra of a heptapeptide, again constrained to the  $\alpha$ -helical geometry. Early observations have shown that heptapeptides of specific sequences are capable of generating spectral responses typical of an  $\alpha$ -helix in nonaqueous solution.<sup>37</sup> Previously,<sup>28</sup> we attempted to model VCD spectra of peptides on the basis of *ab initio* calculations for the dimer molecule  $\text{CH}_3\text{—CONH—CH}_2\text{—CONH—CH}_3$  with the hope that local interactions would dominate the spectral response. Now we extend such a calculation to a model heptapeptide,  $\text{H—(CH}_2\text{—CONH)}_7\text{—CH}_3$ , using our method for transfer of atomic properties. To accomplish this, the “dimer” FF and APT and AAT parameters were taken for transfer onto the heptapeptide geometry. The molecule is too “big” to allow a high-level *ab initio* calculation, particularly considering our local constraints. Thus the geometry of the heptapeptide was initially optimized by energy minimization using the semiempirical PM3 method,<sup>38</sup> as implemented in the MOPAC programs.<sup>39</sup> As for the tripeptide calculations indicated previously, the main peptide chain torsion angles were constrained to mimic an ideal  $\alpha$ -helix ( $\omega = 180^\circ$ ,  $\phi = -47^\circ$ ,  $\psi = -57^\circ$ ). The PM3 force field was then used as the starting point in the transferred FF simulation. In other words, the PM3 long range interactions calculated for the “big” heptapeptide were kept in the final FF when there was not a suitable force constant for transfer from the “small” dimer fragment.

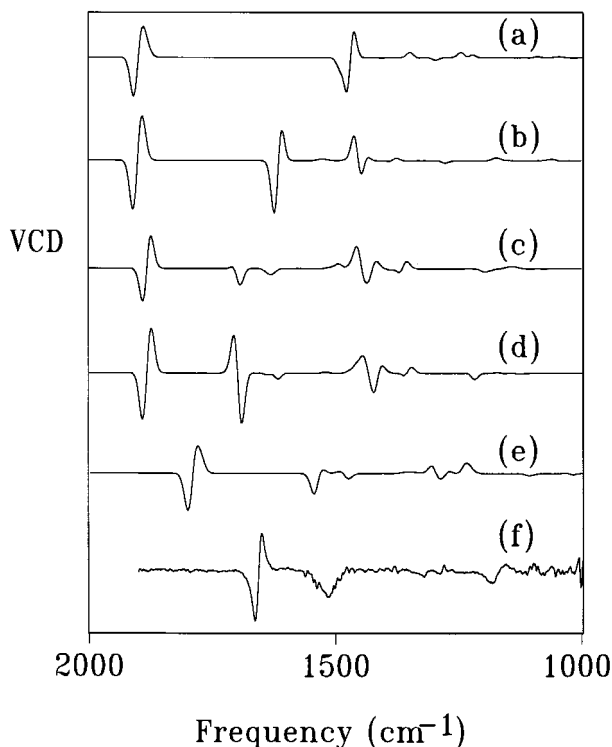
Next, the small molecular fragments used as a source of force constants for the transferred FF were constructed. For this test, in analogy to the above computations, the geometry of the dimer molecule was optimized at the HF/4-31G level, using similar constraints.<sup>28</sup> From our earlier calculations for this dimer molecule,<sup>28</sup> the AATs and APTs at the same HF/4-31G level are available. In addition, a transfer of these tensor properties onto the heptapeptide in terms of internal coordinates was previously constructed.<sup>30</sup> We choose here to use a Cartesian transfer of the dimer (HF/4-31G) results due to the availability of internal coordinate transfer for comparison and due to the well-documented behavior for the dimer when constrained to various stereochemistries. A final “best case result,” obtained by transferring the tripep-

tide force constants obtained at the B3LYP/6-31G\*\* level as described above, was also prepared to investigate the effects of FF improvement. To further exemplify the power of our transfer techniques, Cartesian derivatives of the ROA tensors ( $\alpha$ ,  $G'$ ,  $A$ ) were obtained for the *N*-methylacetamide (NMA) molecule at the HF/6-31G level. These were used as the basis for an even more approximate ROA simulation, which again can be compared to a growing body of ROA data on proteins and peptides.<sup>20</sup> Note that calculation of the ROA tensors is the most time-consuming operation of all the simulations discussed. Then the parameters were transferred from the smaller fragments onto the heptapeptide. An overlap of the structure of the tripeptide and the "big" heptapeptide is shown in Figure 4.

In Figure 5 just the VCD spectra simulated for the heptapeptide are compared for different transfer schemes. For simulation, a uniform bandwidth of  $10\text{ cm}^{-1}$  was used, corresponding to a typical experimental resolution used for VCD of protein samples. As the lowest level of FF calculation against which one could reference the following simulations, spectrum (a) was obtained using the PM3 force field for determining the frequencies, while the intensities were derived from HF/4-31G atomic polar and atomic axial tensors transferred from the dimer using the Cartesian method (the



**FIGURE 4.** An example of an overlap between the model heptapeptide and the tripeptide (marked atoms). A special overlap was defined for each atom pair during the transfer.



**FIGURE 5.** Simulated VCD spectra of the heptapeptide. From top to bottom: (a) the PM3 force field and HF/4-31G APTs and AATs; (b) the diagonal force constants  $[\partial^2 E / (\partial r_\alpha^\lambda \partial r_\beta^\lambda)]$  from an HF/4-31G calculation on the dimer transferred into the FF; (c) the rest of the FF transferred from the dimer; (d) the same transfer done in internal coordinates; (e) the FF and APTs and AATs transferred from the tripeptide B3LYP/6-31G\*\* calculation; (f) an experimental VCD of an  $\alpha$ -helical oligopeptide (Met<sub>2</sub>Leu)<sub>8</sub>.<sup>40</sup>

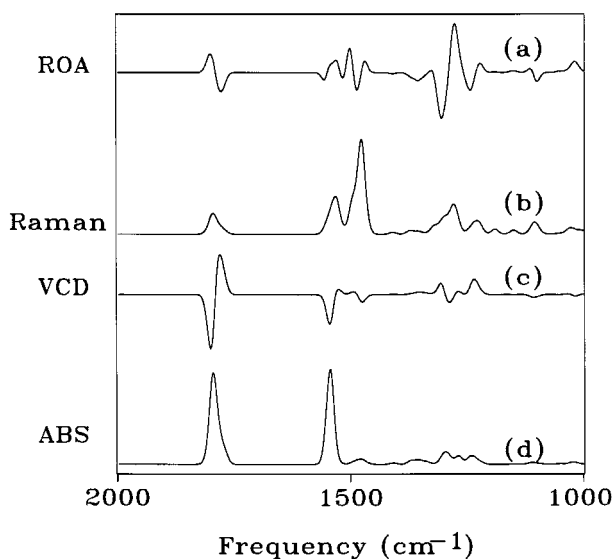
MOPAC program cannot calculate the APT and AAT values). In the PM3 calculation, the amide I–II splitting is severely overestimated, leading to an overlap of amide II with amide III and C–H deformation modes.

For spectrum (b) diagonal  $[\partial^2 E / (\partial r_\alpha^\lambda \partial r_\beta^\lambda)]$  HF force constants and for spectrum (c) also off-diagonal  $[\partial^2 E / (\partial r_\alpha^\mu \partial r_\beta^\lambda), \mu \neq \lambda]$  ones were transferred from the dimer with the Cartesian method and substituted into the PM3 force field. As is apparent in the figure, the substituted off-diagonal force constants significantly influence the resultant spectral pattern, mostly in the amide II region ( $\sim 1550\text{ cm}^{-1}$  experimentally, but overestimated to about  $1700\text{--}1750\text{ cm}^{-1}$  in the HF/4-31G FF). In (c), the amide II and CH<sub>2</sub> deformation bands are resolved resulting in the characteristic  $\alpha$ -helix negative amide II VCD along with the positive couplet amide I VCD seen in all calculations. The amide III

VCD becomes more complex but remains net positive. Internal-coordinate-based transfer of FF and APT and AAT parameters was used to yield spectrum (d).<sup>28</sup> By comparison, spectrum (c), as obtained by the Cartesian transfer, compares much better to the typical experimental VCD spectra of  $\alpha$ -helical peptides, again particularly in the amide II region where only a dominant negative VCD signal is normally seen for  $\alpha$ -helices [see, e.g., spectrum (f)].<sup>40</sup>

Finally, the *ab initio* parameters from the DFT tripeptide were transferred in Cartesian coordinates to yield spectrum (e). Here, we see that the main impact of using the trimer parameters is a sharp improvement in the frequencies for amide I, although amide II is now calculated at too low a frequency. The overall improvement is a consequence of the better, Becke 3LYP/6-31G\*\* force field. In all calculations the characteristically positive couplet VCD for amide I is found. In (e) and (c) the negative amide II and dispersed net positive amide III VCD is also seen. Not shown is the consistent negative couplet VCD for the amide A (N—H stretch) band which also agrees with experiment, but the computed intensities are very weak. The general distortion of the amide II and III bands is partially attributable to our having modeled an oligoglycine molecule. As a consequence, the  $\alpha$ —CH<sub>2</sub> groups give rise to scissoring modes at  $\sim 1450\text{ cm}^{-1}$  that affect amide II yet would not normally be seen in typical peptides, and also to CH<sub>2</sub> rocking and wagging modes that are considerably different from the normal proteinic amino acid tertiary C <sub>$\alpha$</sub> —H deformations known to interact with the amide III modes.<sup>31</sup>

As a last example of parameter transfer the best heptapeptide force field was also used to generate a "complete set" of the most common vibrational spectra, including ROA and Raman (Fig. 6a and b, respectively), using the atomic tensors calculated for *N*-methylacetamide for comparison to experimental ROA spectra of peptides.<sup>40</sup> Obviously, this approach of transferring from NMA totally neglects the local chirality of the  $\alpha$ —carbon site. Because ROA appears to be a highly local phenomenon with the decisive contribution of the  $\alpha$ —C chirality, this becomes somewhat of a calculational exercise at this point. Nevertheless, we can make some comparisons of the spectra to the simulated VCD and absorption spectra (Fig. 6c and d). The main feature we feel that can be learned from these simulations is the relatively weak amide I and II ROA as compared to the amide III ROA and the negative couplet amide I ROA. Indeed,



**FIGURE 6.** Simulated (a) ROA, (b) Raman, (c) VCD, and (d) IR adsorption spectra of the heptapeptide, generated using the combined transfer as described in the text.

such behavior has been confirmed experimentally in numerous protein ROA studies.<sup>41</sup>

## Discussion

We have found no intrinsic disadvantages and have identified some advantages of the Cartesian transfer of FF, APT, and AAT parameters for semi-rigid systems as compared to the more traditional transfer in terms of internal coordinates. Nevertheless, the transfer concept itself cannot be used in many cases; for example, in aromatic systems with conjugated  $\pi$ -bonds the local electron densities of atoms are modified by long range interactions so that atomic-based transfer is destined to be compromised in accuracy. Additionally, there is no simple way of handling flexible molecules whose low frequency modes involve quite nonlocal motions, other than to focus only on the more localized high energy modes and then to sample all possible conformations and appropriately weight their contributions.

A Cartesian transfer of an HF force field scaled in some internal coordinate basis to improve frequency prediction is technically feasible, although part of the simplicity of our method would be lost. It now seems wiser to employ DFT methods (e.g., with the B3LYP parameterization) and to analyze the simulated spectra by correlating the offset be-

tween the theoretical and experimental frequencies, thereby accepting the remaining frequency error as a fundamental limitation of the technique. Any nonuniform scaling can cost a good *ab initio* force field one of its main strengths, the nonparameterized relative ordering of frequencies which is often used to aid in assignment. Comparison of calculated and experimental frequencies without regard to other properties, most generally spectral intensities, can lead to serious errors in larger, more complex molecules, which tend to have near degenerate modes. According to our experience a careful analysis with the *ab initio* force field and intensity distribution is usually more rewarding in terms of understanding the spectra than is an introduction of empirical scale factors.

The assembly of force fields for "bigger" molecules from smaller fragments can substantially decrease the computational time for FF calculation. When the force field for larger molecules is transferred from smaller fragments, we expect the computational effort to scale as the second power of number of atoms. More importantly, it can provide a realistic means of developing FF for "big" molecules with at least some of the positive attributes of the *ab initio* FF now commonly used for "small" molecules. Long range forces can be modeled at a lower level of approximation, because they have a minor influence on observed frequencies, insofar as one can assume they are correct enough to provide a proper mode ordering. A pure *ab initio* approach still does not seem to be reasonable for molecules of the size of oligopeptides and oligonucleotides, which are increasingly the target of biophysical vibrational spectroscopic studies.

## Conclusions

The transfer of molecular force field and other molecular property tensors can be done directly in Cartesian coordinates. The method is easier to implement and numerically more stable than the classical approach using intrinsic coordinates. Furthermore, the force field and the spectral intensity parameters can be transferred simultaneously. The method can be extended for the transfer of higher (anharmonic) derivatives. Finally, we have shown the transfer to be especially convenient for modeling the vibrational optical activity of biopolymers.

## Acknowledgment

This work was supported by the Grant Agency of the Czech Republic (Grant GA/203/95/0105). Work at the UIC was additionally supported by a grant from the National Institutes of Health (GM30147 to T. A. K.).

## References

1. G. Fogarasi and P. Pulay, In *Vibrational Spectra and Structure*, Vol. 14, Elsevier, New York, 1985.
2. T. C. Cheam and S. Krimm, *J. Mol. Struct. (Theochem)*, **206**, 173 (1990).
3. T. A. Keiderling, In *Circular Dichroism and the Conformational Analysis of Biomolecules*, G. D. Fasman, Ed., Plenum, New York, 1996, p. 555.
4. L. D. Barron, L. Hecht, and A. F. Bell, In *Circular Dichroism and the Conformational Analysis of Biomolecules*, G. D. Fasman, Ed., Plenum Press, New York, 1996, p. 653.
5. P. J. Stephens, F. J. Devlin, C. S. Ashvar, C. F. Chabalowski, and N. J. Frisch, *Faraday Disc.*, **99**, 103 (1994).
6. P. L. Polavarapu and Z. Deng, *Faraday Disc.*, **99**, 151 (1994).
7. R. D. Amos, N. C. Handy, K. J. Jalkanen, and P. J. Stephens, *Chem. Phys. Lett.*, **133**, 21 (1987).
8. P. Bouř and T. A. Keiderling, *J. Am. Chem. Soc.*, **114**, 9100 (1992).
9. D. Papoušek and M. R. Aliev, In *Molecular Vibrational-Rotational Spectra*, Academia, Prague, 1982.
10. P. Bouř, *J. Phys. Chem.*, **98**, 8862 (1994).
11. P. Bouř and L. Bednářová, *J. Phys. Chem.*, **99**, 5961 (1995).
12. L. D. Barron, In *Molecular Light Scattering and Optical Activity*, Cambridge University Press, Cambridge, UK, 1982.
13. P. J. Stephens, *J. Phys. Chem.*, **89**, 748 (1985).
14. L. Hecht and L. A. Nafie, *Molec. Phys.*, **72**, 441 (1991).
15. P. J. Stephens, *J. Phys. Chem.*, **91**, 1712 (1987).
16. K. J. Jalkanen, P. J. Stephens, R. D. Amos, and N. C. Handy, *J. Phys. Chem.*, **92**, 1781 (1988).
17. P. L. Polavarapu, P. K. Bose, L. Hecht, and L. D. Barron, *J. Phys. Chem.*, **97**, 11211 (1993).
18. R. D. Amos, *CADPAC, Version 5.0*, SERC Laboratory, Daresbury, UK, 1984, issue 1990.
19. P. L. Polavarapu, *J. Phys. Chem.*, **94**, 8106 (1990).
20. C. N. Tam, P. Bour, and T. A. Keiderling, *J. Phys. Chem.* (in press).
21. W. D. Allen and A. G. Császár, *J. Chem. Phys.*, **98**, 2983 (1993).
22. B. G. Johnson, P. M. W. Gill, and J. A. Pople, *J. Chem. Phys.*, **98**, 5612 (1993).
23. M. Ehrenberg, *Acta Crystallogr.*, **19**, 698 (1965).
24. M. J. Frisch, et al., *GAUSSIAN 94*, Gaussian Inc., Pittsburgh, PA, 1994.
25. T. L. Blundell and L. N. Johnson, In *Protein Crystallography*, Academic Press, New York, 1976.
26. B. Schneider, A. Hoření, H. Pivcová, and J. Honzl, *Collect. Czech. Chem. Commun.*, **30**, 2196 (1964).

27. P. L. Polavarapu, Z. Debg, and C. S. Ewig, *J. Phys. Chem.*, **98**, 9919 (1994).
28. P. Bouř and T. A. Keiderling, *J. Am. Chem. Soc.*, **115**, 9602 (1993).
29. K. L. Bak, F. J. Devlin, C. S. Ashvar, P. R. Taylor, M. J. Frisch, and P. J. Stephens, *J. Phys. Chem.*, **99**, 14918 (1995).
30. P. Bouř, PhD. thesis, Academy of Sciences, Prague, 1992.
31. M. Diem, M. R. Oboodi and C. Alva, *Biopolymers*, **23**, 1917 (1984).
32. T. B. Freedman, L. A. Nafie, and T. A. Keiderling, *Biopolymers (J. Pept. Sci.)*, **37**, 265 (1995).
33. R. D. Singh, and T. A. Keiderling, *Biopolymers*, **26**, 237 (1981).
34. B. B. Lal and L. A. Nafie, *Biopolymers*, **21**, 2161 (1982).
35. A. C. Sen and T. A. Keiderling, *Biopolymers*, **23**, 1519 (1984).
36. P. Maloň, R. Kobrinskaya, and T. A. Keiderling, *Biopolymers*, **27**, 733 (1988).
37. C. Toniolo, *CRC Crit. Rev. Biochem.*, **9**, 1 (1980).
38. J. J. P. Stewart, *J. Comput. Chem.*, **10**, 209 (1989).
39. *MOPAC, Version 6.00*, Frank J. Seiler Research Lab, U.S. Air Force Academy.
40. S. C. Yasui, T. A. Keiderling, and R. Katachai, *Biopolymers*, **26**, 1407 (1987).
41. S. J. Ford, Z. Q. Wen, L. Hecht, and L. D. Barron, *Biopolymers*, **34**, 303 (1994).